Uncertainty quantification and reduction for the monotonicity properties of expensive-to-evaluate computer models

Julien Bect§, Nicolas Bousquet†, Bertrand Iooss†, Shijie Liu§, Alice Mabille§,
Anne-Laure Popelin†, Thibault Rivière§, Rémi Stroh§, Roman Sueur†, Emmanuel Vazquez§

§ SUPELEC, Gif-sur-Yvette, France   † EDF R&D, Chatou, France

Abstract. We consider the problem of estimating monotonicity properties of a scalar-valued numerical model—e.g., a finite element model combined with some post-processing. Several quantitative monotonicity indicators are introduced. Since the evaluation of the numerical model is usually time-consuming, these indicators have to be estimated with a small budget of evaluations.

We adopt a Bayesian approach, where the numerical model itself is modeled as a Gaussian process. We estimate the monotonicity indicators, and quantify the uncertainty surrounding them, through conditional simulations of the Gaussian process partial derivatives. The approach is illustrated with a numerical model of a passive component in a power plant.

Future work will leverage this framework together with the Stepwise Uncertainty Reduction principle to create sequential design strategies, in order get an improved knowledge of the monotonicity properties of the model.

Understanding the structure of numerical models

Consider a (deterministic) numerical model:

with the following assumptions :

- $x \in \mathbb{X} \subset \mathbb{R}^d$,
- $f(x) \in \mathbb{R}$ (scalar output),
- f is expensive to evaluate.

Computer experiments, when properly designed, are a useful tool to discover (or confirm) structural properties of such a numerical model:

- active/inactive variables (screening)
- additive responses, low-order interactions (S.A.),
- and, in this work: monotonicity properties

using a limited number of evaluations (runs of the code).

Partial monotonicity properties

We assume that

- $\mathbb{X}$ is an hyper-rectangle: $\mathbb{X} = \prod_{j=1}^{d} [a_j; b_j]$,
- f admits (at least first-order) partial derivatives,
- but ∇f is not available.

Definition (increasing case). f is said to be increasing with respect to its j^{th} variable if

$$\forall x^{(-j)} \in \prod_{k \neq j} [a_k; b_k], \quad x^{(j)} \mapsto f(x) \text{ is increasing.}$$

Proposition. f is increasing with respect to its j^{th} variable if, and only if,

$$\frac{\partial f}{\partial x^{(j)}}(x) \geq 0 \qquad \forall x \in \mathbb{X}.$$

Quantitative monotonicity indicators

Several quantities of interest (nonlinear functionals of f):

- extrema of the partial derivatives

$$M_j^-(f) = \min_{x \in \mathbb{X}} \frac{\partial f}{\partial x^{(j)}}(x), \quad \text{and } M_j^+(f) = \max_{x \in \mathbb{X}} \frac{\partial f}{\partial x^{(j)}}(x),$$

- "positivity rate" of the partial derivatives

$$\alpha_j(f) = \mu\left(\Gamma_j(f)\right), \quad \text{where } \Gamma_j(f) = \left\{\partial f / \partial x^{(j)} \geq 0\right\},$$

where μ is a given probability measure on $\mathbb{X}$.

Proposed (Bayesian) approach:

- Endow f with a Gaussian Process prior,
- Estimate the monotonicity indicators, and quantify the uncertainty surrounding them, using conditional simulations of the partial derivatives.

Prediction and simulation of derivatives: cokriging

Consider a classical GP model:

$$f \mid \beta, \theta \sim \mathcal{GP}\left(\sum_{j=1}^{\ell} \beta_j h_j(\cdot),\ k_\theta(\cdot, \cdot)\right)$$

- $h_1, \ldots, h_\ell$ known functions, which admit first-order partial derivatives (typically, polynomial functions)
- k_θ a stationary covariance function (for simplicity).

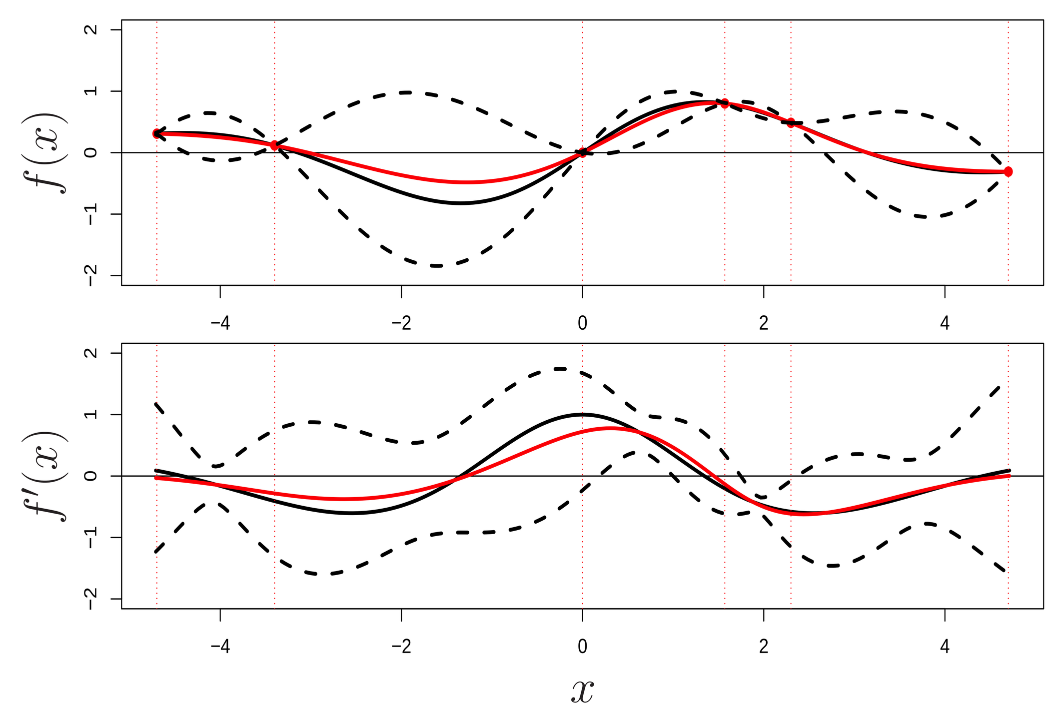
Theorem (see, e.g., Stein, 1999). Let $\tilde{k}_\theta(h) = k_\theta(x, x+h)$. The random process f is mean-square differentiable if, and only if, $\tilde{k}_\theta$ is twice differentiable at $h = 0$.

Theorem (Scheuerer, 2010). The random process f is mean-square differentiable if, and only if, the partial derivatives exist almost surely in Sobolev's weak sense.

Useful fact. Cokriging is just a special case of kriging with an auxiliary discrete variable:

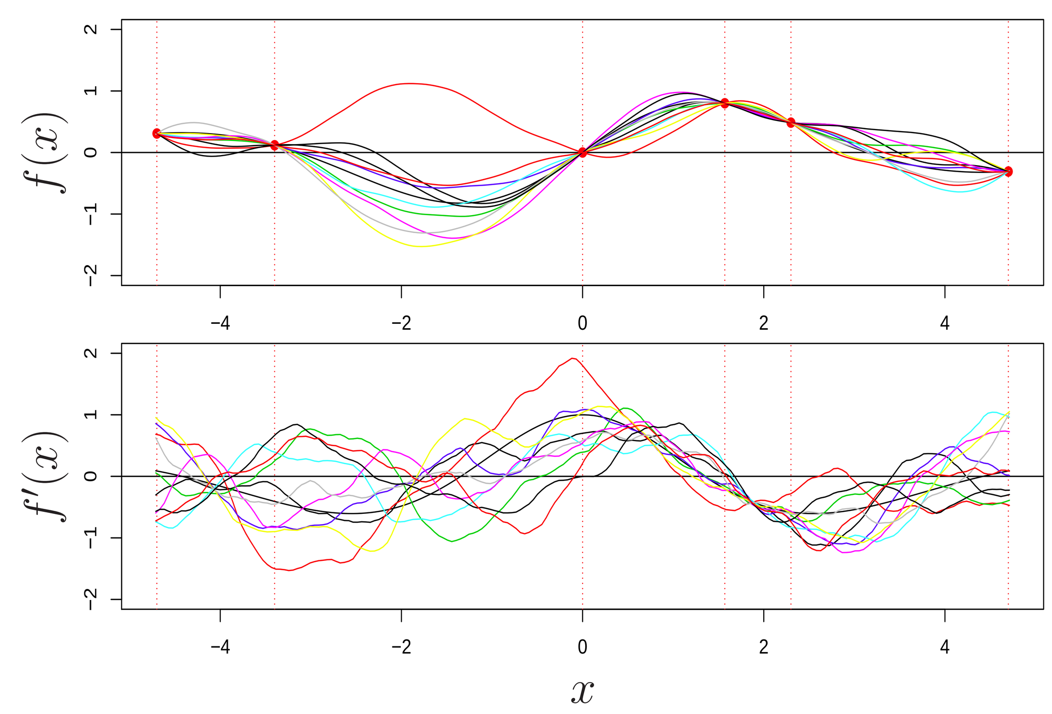
$$\tilde{f}(x, j) = \begin{cases} f(x) & \text{if } j = 0, \\ \partial f / \partial x^{(j)} & \text{if } j > 0. \end{cases}$$

- No need to replace your favorite kriging software (if it is flexible enough...)



Cokriging example: prediction and 95%-CI

The true function f (black line) is observed at six locations (red vertical dotted lines). The kriging prediction (red line) and CI (black dashed lines) are performed using a zero-mean GP model, with a Matérn 5/2 covariance function.



Cokriging example: conditional simulations

Same setup as above. Observe that the samplepaths of f' are rougher than those of f: indeed, f has a Matérn 5/2 covariance function, while f' has a covariance function with the same regularity as a Matérn 3/2 covariance function.

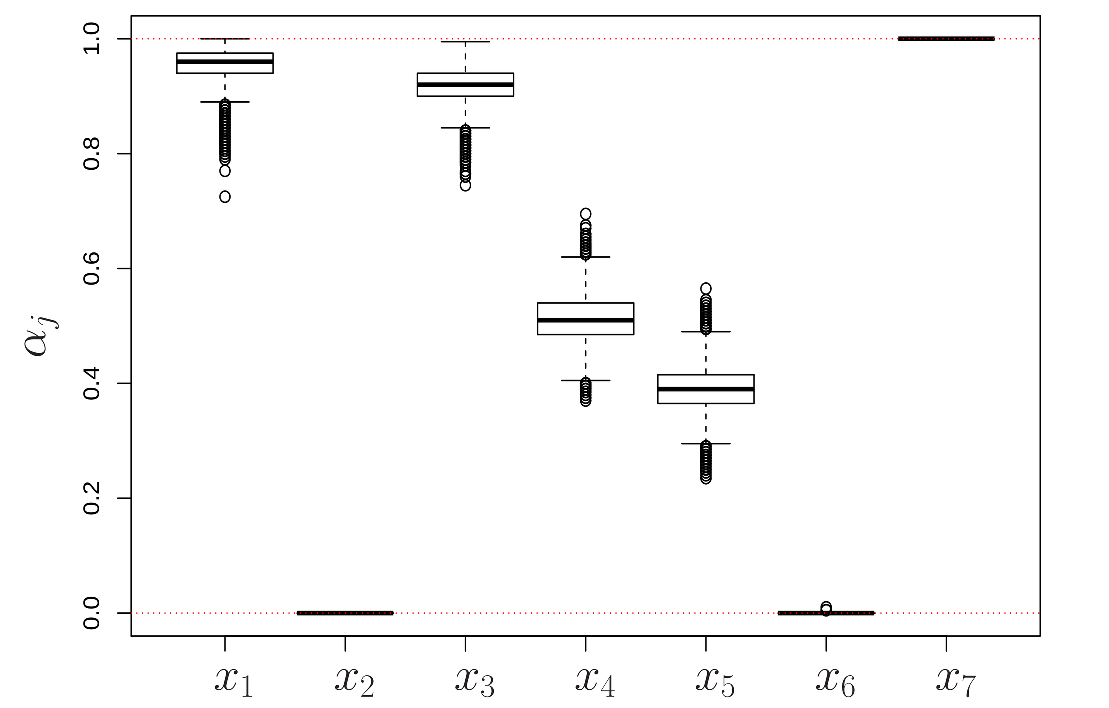
Industrial test case ($d = 7$)

The methodology has been applied to an industrial test case proposed by EDF R&D

- Goal: assess the performance of a passive component in a power plant
- Thermomechanical numerical code, $d = 7$ quantitative input factors.

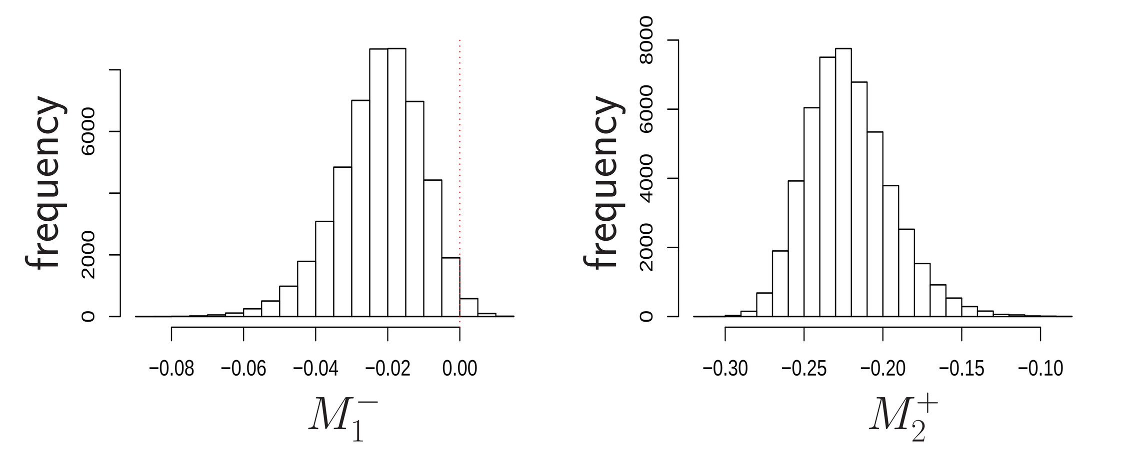
Implementation ⇒ discretizations

- Intractable posterior distributions for the quantitative monotonicity indicators: we rely on conditional simulations,
- Approximate computation of the indicators themselves using a Monte Carlo sample $x_1, \ldots, x_m$ ($m = 200$)
- Simulation scheme
  - $R = 100$ iid draws of the spatial MC m-sample,
  - $S = 500$ iid conditional simulations on each of them,
  - $RS = 50\,000$ (dependent) conditional samplepaths.



Posterior distribution of $\alpha_1, \ldots, \alpha_d$

"Space-filling" design of size $n = 70$. Gaussian Process model: affine mean function, Matérn 5/2 covariance function (similar results obtained with other covariance functions) with MLE parameter estimation. Results in agreement with the prior judgment of an EDF expert of this numerical model.



Posterior distribution of M_1^- and M_2^+

Histrograms based on the same set of $RS = 50\,000$ conditional simulation as above. Left: confirms that, with high credibility, the function is not increasing w.r.t. $x^{(1)}$ on its entire range of definition. Right: the function *might be* decreasing w.r.t. $x^{(2)}$.

Future work

Work is in progress on

- Advanced simulation techniques for excursion sets (see Ginsbourger et al, 2014, for a preview): will replace the crude MC-based technique used here,
- Stepwise uncertainty reduction (using ideas from Bect et al, 2012): will be used to enrich a given initial design to learn more precisely the value of the indicators α_j, M_j^+ and M_j^-, or the set Γ_j itself !

References

J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3):773–793, 2012.

D. Ginsbourger, C. Chevalier, J. Bect, D. Azzimonti, and I. Molchanov. Quantifying and reducing uncertainties on excursion sets under a gaussian random field prior. In *2nd Conf. of the ISNPS, June 12-16, Cadiz, Spain*, 2014.

M. Scheuerer. Regularity of the sample paths of a general second order random field. *Stochastic Processes and their Applications*, 120(10): 1879–1897, 2010.

M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.